

Distinct spreads in vector spaces over finite fields

Ben Lund

Thang Pham

Le Anh Vinh

Abstract

In this short note, we study the distribution of spreads in a point set $\mathcal{P} \subseteq \mathbb{F}_q^d$, which are analogous to angles in Euclidean space. More precisely, we prove that, for any $\varepsilon > 0$, if $|\mathcal{P}| \geq (1 + \varepsilon)q^{\lceil d/2 \rceil}$, then \mathcal{P} generates a positive proportion of all spreads. We show that these results are tight, in the sense that there exist sets $\mathcal{P} \subset \mathbb{F}_q^d$ of size $|\mathcal{P}| = q^{\lceil d/2 \rceil}$ that determine at most one spread.

1 Introduction

Let $q = p^r$ be a large odd prime power, and \mathbb{F}_q be a finite field of order q . For any two points \mathbf{x} and \mathbf{y} in \mathbb{F}_q^d , we define the distance between \mathbf{x} and \mathbf{y} by

$$\|\mathbf{x} - \mathbf{y}\| = (x_1 - y_1)^2 + \cdots + (x_d - y_d)^2.$$

Although it is not a norm, the function $\|\mathbf{x} - \mathbf{y}\|$ has properties similar to the Euclidean norm (for example, it is invariant under orthogonal matrices and translations).

The Erdős-Falconer distance problem asks for the minimum exponent α such that for any set $\mathcal{P} \subseteq \mathbb{F}_q^d$ with $|\mathcal{P}| \gg q^\alpha$, the number of distinct distances determined by \mathcal{P} is at least cq for some positive constant c .¹ Bourgain, Katz, and Tao [1] considered a similar problem on the number of distinct distances determined by a set of points in \mathbb{F}_q^2 . Iosevich and Rudnev [11] proved that for any $\mathcal{P} \subseteq \mathbb{F}_q^d$, if $|\mathcal{P}| \geq 2q^{\frac{d+1}{2}}$, then all distances are determined by \mathcal{P} . The authors of [7] indicated that the exponent $(d+1)/2$ is the best possible in odd dimensions; the correct exponent is not known in even dimensions.

Let S_1 be the unit sphere in \mathbb{F}_q^d , i.e. the set of points $\mathbf{x} \in \mathbb{F}_q^d$ with $\|\mathbf{x}\| = 1$. If \mathcal{P} is a subset in the unit sphere S_1 , the authors of [7] proved that if $|\mathcal{P}| \geq Cq^{\frac{d}{2}}$ for some sufficiently large positive constant C , then there exists $c > 0$ such that the number of distinct distances determined by points in \mathcal{P} is at least cq . We note here that their result even can be stated in a stronger form which will be useful for our later applications. The interested reader can find more details in [7, page 15].

Theorem 1 (Hart et al., [7]). *For $\mathcal{P} \subseteq S_1$ in \mathbb{F}_q^d with $d \geq 3$. Suppose that $|\mathcal{P}| \geq Cq^{\frac{d}{2}}$ for some positive constant C , then the number of distinct distances determined by points in \mathcal{P} is at least $\min\{q/2, Cq/4\}$.*

¹Here and throughout, $X \gg Y$ means that there exists $C > 0$ such that $X \geq CY$.

In 2014, Bennett, Iosevich and Pakianathan [4] studied a generalization of the Erdős distinct distance problem in vector spaces over finite fields. More precisely, they dealt with the distribution of classes of triangles in a point set in \mathbb{F}_q^d . They proved that for any $\mathcal{P} \subseteq \mathbb{F}_q^2$, if $|\mathcal{P}| \gg q^{7/4}$ then \mathcal{P} determines at least a positive proportion of all congruence classes of triangles, where two triangles, denoted by $\Delta(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ and $\Delta(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$, are in the same congruence class if there exist an orthogonal matrix M and a vector $\mathbf{z} \in \mathbb{F}_q^2$ such that $M \cdot \mathbf{a}_i + \mathbf{z} = \mathbf{b}_i$ for $1 \leq i \leq 3$. The threshold $q^{7/4}$ was improved recently to $q^{8/5}$ by Bennett et al. [3] by using a clever combination of Fourier analytic techniques and elementary results from group action theory. The authors of [3] also gave a construction of a point set $\mathcal{P} = \mathcal{A} \times \mathcal{B}$ with $|\mathcal{A}| = q^{1/2-\varepsilon'}$ and $|\mathcal{B}| = q$, and \mathcal{P} determines at most $cq^{3-\varepsilon''}$ for $\varepsilon'' > 0$. Iosevich [10] conjectured that for any $\mathcal{P} \subset \mathbb{F}_q^2$, if $|\mathcal{P}| \geq Cq^{3/2}$ for a sufficiently large constant C then \mathcal{P} determines a positive proportion of all congruence classes of triangles. The interested reader can find more discussions and related problems in [4, 3, 5, 13]. There is also a series of papers dealing with similar problems, see for example [2, 3, 4, 6, 9, 12, 14, 15].

In this paper, we study a similar problem on the number of distinct *spreads* generated by a point set in \mathbb{F}_q^d .

For three points $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{F}_q^d$, the spread between two vectors $\overrightarrow{\mathbf{ab}}$ and $\overrightarrow{\mathbf{ac}}$ in \mathbb{F}_q^d , which is denoted by $S(\overrightarrow{\mathbf{ab}}, \overrightarrow{\mathbf{ac}})$ (or $S(\mathbf{b}, \mathbf{a}, \mathbf{c})$ for simplicity), is defined by

$$S(\overrightarrow{\mathbf{ab}}, \overrightarrow{\mathbf{ac}}) = 1 - \frac{(\overrightarrow{\mathbf{ab}} \cdot \overrightarrow{\mathbf{ac}})^2}{\|\overrightarrow{\mathbf{ab}}\| \cdot \|\overrightarrow{\mathbf{ac}}\|},$$

where $\|\overrightarrow{\mathbf{x}}\| = x_1^2 + \cdots + x_d^2$. If either term in the denominator is 0, then $S(\overrightarrow{\mathbf{ab}}, \overrightarrow{\mathbf{ac}})$ is undefined.

It is clear that this definition is consistent with the square of the sine of the angle between two vectors $\overrightarrow{\mathbf{ab}}$ and $\overrightarrow{\mathbf{ac}}$ in Euclidean space

$$\sin(\theta)^2 = 1 - \frac{(\overrightarrow{\mathbf{ab}} \cdot \overrightarrow{\mathbf{ac}})^2}{\|\overrightarrow{\mathbf{ab}}\| \cdot \|\overrightarrow{\mathbf{ac}}\|}.$$

The following are some properties of the spread between two vectors $\overrightarrow{\mathbf{ab}}$ and $\overrightarrow{\mathbf{ac}}$:

- (i) $S(\overrightarrow{\mathbf{ab}}, \overrightarrow{\mathbf{ac}}) = S(r(\overrightarrow{\mathbf{ab}}), s(\overrightarrow{\mathbf{ab}}))$ for any $r, s \in \mathbb{F}_q^*$,
- (ii) $S(\overrightarrow{\mathbf{ab}}, \overrightarrow{\mathbf{ac}}) = S(\overrightarrow{\mathbf{ac}}, \overrightarrow{\mathbf{ab}})$,
- (iii) $S(\overrightarrow{\mathbf{ab}}, \overrightarrow{\mathbf{ac}}) = S(M \cdot \overrightarrow{\mathbf{ab}}, M \cdot \overrightarrow{\mathbf{ac}})$, where M is an orthogonal matrix.

In 2015, Bennett [2] made the first investigation on the number of distinct spreads determined by points in $\mathcal{P} \subseteq \mathbb{F}_q^d$. In particular, he obtained the following.

Theorem 2 (Theorem 6.5, [2]). *Let \mathcal{P} be a set of points in \mathbb{F}_q^2 . If $|\mathcal{P}| \geq 2q - 1$ then the number of distinct spreads generated by points in \mathcal{P} is q .*

It is clear that Theorem 2 is sharp up to the coefficient of q , since the number of spreads spanned by points in a line of q points is at most one. For higher dimensional cases, Bennett [2] had an observation on a connection between spreads and distances:

A connection between spreads and distances on a sphere: Suppose \mathcal{P}_1 is a subset in the unit sphere S_1 , it is easily to check that $S(\vec{0a}, \vec{0b}) = S(\vec{0c}, \vec{0d})$ with $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathcal{P}_1$ if and only if either $\|\mathbf{a} - \mathbf{b}\| = \|\mathbf{c} - \mathbf{d}\|$ or $\|\mathbf{a} - \mathbf{b}\| = \|\mathbf{c} + \mathbf{d}\|$. Thus if \mathcal{P}_1 determines a positive proportion of all distances then \mathcal{P}_1 generates a positive proportion of all spreads. Therefore if we have a set $\mathcal{P} \subset \mathbb{F}_q^d$ satisfying $|\mathcal{P}| \gg q^{\frac{d+2}{2}}$ then there exists a sphere of radius $t \neq 0$ such that $|S_t \cap \mathcal{P}| \gg q^{d/2}$. From the first property of spread, we may assume that S_t is the unit sphere. It follows from Theorem 1 that $S_1 \cap \mathcal{P}$ determines a positive proportion of all distances, therefore $S_1 \cap \mathcal{P}$ generates a positive proportion of all spreads. In other words, we have proved the following.

Theorem 3 (Theorem 6.3, [2]). *Let \mathcal{P} be a set of points in \mathbb{F}_q^d , with $d \geq 3$. If $|\mathcal{P}| \gg q^{(d+2)/2}$ then \mathcal{P} generates a positive proportion of all spreads.*

We remark here that if \mathcal{P} is a subset in the unit sphere S_1 , the third listed author [14] showed that for $\mathcal{P} \subseteq \mathbb{F}_q^3$ with $|\mathcal{P}| \gg q^{3/2}$, the number of occurrences of a fixed spread γ among \mathcal{P} is $\Theta\left(\frac{|\mathcal{P}|^2}{q}\right)$ if $1 - \gamma$ is not a square in \mathbb{F}_q .

The main purpose of this short note is to give sharp results on the number of distinct spreads generated by a large set in \mathbb{F}_q^d .

Statement of main results: Our first result gives us the number of distinct spreads generated by $\mathcal{P} \subseteq \mathbb{F}_q^d$ with d even.

Theorem 4. *For any $\varepsilon > 0$, there exists $c > 0$ such that the following holds. Let \mathcal{P} be a set of points in \mathbb{F}_q^d with $d \geq 2$ even. If $|\mathcal{P}| \geq (1 + \varepsilon)q^{d/2}$, then the number of distinct spreads determined by \mathcal{P} is at least cq .*

If \mathcal{P} be a subset in \mathbb{F}_q^d with d odd, then we can embed \mathcal{P} in \mathbb{F}_q^{d+1} with the last coordinate of 0. Therefore, as a direct consequence of Theorem 4, we obtain the following result.

Theorem 5. *For any $\varepsilon > 0$, there exists $c > 0$ such that the following holds. Let \mathcal{P} be a set of points in \mathbb{F}_q^d with $d \geq 3$ odd. If $|\mathcal{P}| \geq (1 + \varepsilon)q^{(d+1)/2}$, then the number of distinct spreads determined by \mathcal{P} is at least cq .*

Sharpness of results: We show that the conditions on the size of \mathcal{P} in Theorem 4 and Theorem 5 are sharp.

Theorem 6. *Let \mathbb{F}_q be a finite field of order q with $q \equiv 1 \pmod{4}$. Then there exists a subset \mathcal{P} in \mathbb{F}_q^d with $d \geq 4$ even such that $|\mathcal{P}| = q^{d/2}$ and there is no spread determined by points in \mathcal{P} .*

Theorem 7. *Let \mathbb{F}_q be a finite field of order q with $q \equiv 1 \pmod{4}$. Then there exists a subset \mathcal{P} in \mathbb{F}_q^d with $d \geq 3$ odd such that $|\mathcal{P}| = q^{(d+1)/2}$ and the number of distinct spreads determined by points in \mathcal{P} is at most one.*

The rest of this note is organized as follows: in Section 2 we give a proof of Theorem 4, in Section 3, we give proofs of Theorems 6 and 7.

2 Proof of Theorem 4

To prove Theorem 4, we make use of the following theorem due to the first listed author and Saraf in [8].

Theorem 8 (Corollary 5, [8]). *For any $\varepsilon > 0$ and $\mathcal{P} \subseteq \mathbb{F}_q^d$ with $|\mathcal{P}| \geq (1 + \varepsilon)q^{d-1}$, the number of lines spanned by \mathcal{P} is bounded below by $\alpha_\varepsilon q^{2d-2}$, where $\alpha_\varepsilon = \varepsilon^2(1 + \varepsilon + \varepsilon^2)^{-1}$.*

By using Theorem 8, we are able to show in our following theorem that if the cardinality of \mathcal{P} is much smaller than q^{d-1} , we still have many distinct lines spanned by \mathcal{P} .

Theorem 9. *For any $0 < \varepsilon < q-1$, let $\mathcal{P} \subseteq \mathbb{F}_q^d$ with $|\mathcal{P}| \geq (1 + \varepsilon)q^{k-1}$. Then, the number of lines spanned by \mathcal{P} is bounded below by $(1 - o(1))\alpha_\varepsilon q^{2k-2}$.*

Proof. Assume that $(1 + \varepsilon)|\mathcal{P}|$ is an integer, and remove all but exactly $(1 + \varepsilon)|\mathcal{P}|$ points from \mathcal{P} . Error introduced by assuming that $(1 + \varepsilon)|\mathcal{P}|$ is an integer will only affect the $o(1)$ term in the result, and removing points from \mathcal{P} only decreases the number of lines spanned by \mathcal{P} .

Let π' be a uniformly random projection from \mathbb{F}_q^d to \mathbb{F}_q^k .

Let \mathbf{a}, \mathbf{b} be two arbitrary distinct points in \mathbb{F}_q^d . We claim that the probability that $\pi'(\mathbf{a}) = \pi'(\mathbf{b})$ is less than q^{-k} . Note that, if $\pi'(\mathbf{a}) = \pi'(\mathbf{b})$, then $\pi'(\mathbf{a} - \mathbf{x}) = \pi'(\mathbf{b} - \mathbf{x})$ for an arbitrary translation vector \mathbf{x} . Hence, we may without loss of generality assume that $\mathbf{a} = \mathbf{0}$. Then, the question of whether $\pi'(\mathbf{a}) = \pi'(\mathbf{b})$ reduces to the question of whether \mathbf{b} lies in the kernel of π' , which is a uniformly random $(d - k)$ -dimensional linear subspace. This probability is $(q^{d-k} - 1)/q^d < q^{-k}$.

Hence, by linearity of expectation, the expected number of pairs $\mathbf{a}, \mathbf{b} \in \mathcal{P}$ such that $\pi'(\mathbf{a}) = \pi'(\mathbf{b})$, denoted by E_{coll} , is $E_{\text{coll}} < \binom{|\mathcal{P}|}{2} q^{-k} = (1 - o(1))(1 + \varepsilon)^2 q^{k-2}/2$. In particular, there exists a projection π from \mathbb{F}_q^d to \mathbb{F}_q^k such that the number of such collisions is at most E_{coll} . By a Bonferroni inequality, the image $\pi(\mathcal{P})$ of \mathcal{P} has size at least $|\pi(\mathcal{P})| \geq |\mathcal{P}| - E_{\text{coll}}$. Thus $|\pi(\mathcal{P})| = (1 - o(1))|\mathcal{P}|$. The conclusion of the theorem follows from Theorem 8, and the observation that $\pi(\mathcal{P})$ does not span more lines than \mathcal{P} . \square

Corollary 10. *Let \mathcal{P} be a set of points in \mathbb{F}_q^d with d even, and \mathcal{L} be the set of spanned lines by \mathcal{P} . Suppose that $|\mathcal{P}| = (1 + \varepsilon)q^{d/2}$, $\varepsilon > 0$, then there exists a point \mathbf{p} in \mathcal{P} such that it is incident to at least $(1 - o(1))\frac{\alpha_\varepsilon}{1 + \varepsilon}q^{d/2}$ lines from \mathcal{L} .*

Proof. It follows from Theorem 9 that the number of lines spanned by \mathcal{P} is bounded below by $(1 - o(1))\alpha_\varepsilon q^d$. By the pigeonhole-principle, there exists a point \mathbf{p} in \mathcal{P} such that it is incident to at least $(1 - o(1))\frac{\alpha_\varepsilon}{1+\varepsilon}q^{d/2}$ lines, and the corollary follows. \square

Proof of Theorem 4: By Corollary 10, if $|\mathcal{P}| \geq (1 + \varepsilon)q^{d/2}$, then there exists a point \mathbf{p} in \mathcal{P} such that it is incident to at least $cq^{d/2}$ lines that are spanned by \mathcal{P} for some positive constant c depending on ε .

Suppose $d = 2$. Then, if $\sqrt{-1} \in \mathbb{F}_q$, then there are $q - 1$ points of \mathbb{F}_q^2 at distance 0 from \mathbf{p} , lying on a single isotropic line with slope $\sqrt{-1}$. If $\sqrt{-1} \notin \mathbb{F}_q$, then there is no point distinct from \mathbf{p} at zero distance from \mathbf{p} . If $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{P}$ such are in three distinct, non-isotropic lines incident to \mathbf{p} , then an easy calculation shows that $S(\mathbf{a}, \mathbf{p}, \mathbf{b}) \neq S(\mathbf{a}, \mathbf{p}, \mathbf{c})$, which proves Theorem 4 in the case $d = 2$.

Suppose $d > 2$. We denote the set of lines incident to \mathbf{p} by \mathcal{L}' . One can check that there exists a sphere S_t of radius $t \neq 0$ such that $|S_t \cap \mathcal{L}'| \geq \frac{cq^{d/2}}{2}$. Without loss of generality, we assume that $\mathbf{p} = \mathbf{0}$ and $t = 1$. Theorem 1 implies that $S_1 \cap \mathcal{L}'$ determines a positive proportion of all distances. Thus Theorem 4 follows from the connection between spreads and distances given in the introduction. \square

3 Proofs of Theorems 6 and 7

In this section, we will use the construction given in [7, Lemma 5.1]. We denote $i = \sqrt{-1}$, which is guaranteed to exist since we assume that $q \equiv 1 \pmod{4}$.

Proof of Theorem 6: Suppose $d = 2m$ with $m \geq 2$. Let \mathcal{P} be the subspace spanned by $\mathbf{v}_1, \dots, \mathbf{v}_m$, where

$$\mathbf{v}_1 = (1, i, 0, \dots, 0), \mathbf{v}_2 = (0, 0, 1, i, 0, \dots, 0), \dots, \mathbf{v}_m = (0, \dots, 0, 1, i).$$

It is easy to check that all vectors \mathbf{v}_i are null orthogonal, i.e. $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $1 \leq i, j \leq m$. Since $\|\mathbf{v}_i\| = 0$ for all $1 \leq i \leq m$, it follows from the definition of spread that there is no spread determined by three vectors in \mathcal{P} . On the other hand, the size of \mathcal{P} is $q^{d/2}$, which ends the proof of the theorem. \square

Proof of Theorem 7: Suppose $d = 2m + 1$ with $m \geq 2$. Let \mathcal{P} be the subspace spanned by $\mathbf{v}_1, \dots, \mathbf{v}_{m+1}$, where

$$\mathbf{v}_1 = (1, i, 0, \dots, 0), \mathbf{v}_2 = (0, 0, 1, i, 0, \dots, 0), \dots, \mathbf{v}_m = (0, \dots, 0, 1, i), \mathbf{v}_{m+1} = (0, \dots, 0, 1).$$

We have the size of \mathcal{P} is $q^{(d+1)/2}$. It is easy to check that the spread spanned by any triple of points in \mathcal{P} is either undefined or one. Thus the number of distinct spreads spanned by \mathcal{P} is at most one. This concludes the proof of the theorem. \square

4 Acknowledgments

Research of the first listed author was supported by NSF grant CCF-1350572. The second listed author was partially supported by Swiss National Science Foundation grants 200020-162884 and 200020-144531. The research of the third listed author is funded by the National Foundation for Science and Technology Development Project. 101.99-2013.21.

References

- [1] J. Bourgain, N. Katz, T. Tao, *A sum-product estimate in finite fields and applications*, Geom. Funct. Anal. **14** (2004), 27–57.
- [2] M. Bennett, *Right Angles in \mathbb{F}_q^d* , arXiv:1511.08942 (2015).
- [3] M. Bennett, D. Hart, A. Iosevich, J. Pakianathan, M. Rudnev, *Group actions and geometric combinatorics in \mathbb{F}_q^d* , to appear in Forum Mathematicum 2016.
- [4] M. Bennett, A. Iosevich, and J. Pakianathan, *Three-point configurations determined by subsets of \mathbb{F}_q^2 via the Elekes-Sharir Paradigm*, Combinatorica **34**(6) (2014): 689–706.
- [5] D. Covert, D. Hart, A. Iosevich, S. Senger, I. Uriarte-Tuero, *A FurstenbergKatznelsonWeiss type theorem on $(d + 1)$ -point configurations in sets of positive density in finite field geometries*, Discrete Mathematics, **311**(6) (2011), 423–430.
- [6] B. Hanson, B. Lund, O. Roche-Newton, *On distinct perpendicular bisectors and pinned distances in finite fields*, Finite Fields and Their Applications, **37** (2016), 240–264.
- [7] D. Hart, A. Iosevich, D. Koh, M. Rudnev, *Averages over hyperplanes, sum-product theory in vector spaces over finite fields and the Erdős–Falconer distance conjecture*, Transactions of the American Mathematical Society, **363**(6) (2011), 3255–3275.
- [8] B. Lund, S. Saraf, *Incidence bounds for block designs*, SIAM Journal on Discrete Mathematics, **30**(4) (2016), 1997–2010.
- [9] A. Iosevich, M. Rudnev, Y. Zhai, *Areas of triangles and Becks theorem in planes over finite fields*, Combinatorica, 2012, 1–14.
- [10] A. Iosevich, *Personal communication*, 2016.
- [11] A. Iosevich and M. Rudnev, *Erdős distance problem in vector spaces over finite fields*, Trans. Amer. Math. Soc. **359** (2007), 6127–6142.
- [12] I. Shparlinski, *On point sets in vector spaces over finite fields that determine only acute angle triangles*, Bulletin of the Australian Mathematical Society **81**(01) (2010): 114–120.

- [13] H. Pham, T. Pham, L. A. Vinh, *An improvement on the number of simplices in \mathbb{F}_q^d* , submitted, arxiv: 1608.06398v1.
- [14] L. A. Vinh, *The number of occurrences of a fixed spread among n directions in vector spaces over finite fields*, Graphs Combin. **29** (2008), no. 6, 1943–1949.
- [15] L. A. Vinh, *The Erdős-Facolner distance problem in subsets of spheres over finite fields*, SIAM Journal on Discrete Mathematics, **25**(2) 681–684 (2011).

Department of Computer Science,
 Rutgers, The State University of New Jersey, NJ
 E-mail: lund.ben@gmail.com

Department of Mathematics,
 EPF Lausanne
 Switzerland
 E-mail: thang.pham@epfl.ch

University of Education,
 Vietnam National University
 Viet Nam
 E-mail: vinhla@vnu.edu.vn